Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 79 (2016) 579 – 588

Procedia Computer Science

7th International Conference on Communication, Computing and Virtualization 2016

Truncated DCT and Decomposed DWT SVD features for Image Retrieval

NST Sai[a,1] ,Ravindra Patil*[b,2] , Shailesh Sangle[b,3,]Bhushan Nemade[d]

[a]TechMahinda, Mumbai -400121, Maharashtra, India,[b] Thakur College of Engg. and Tech., Mumbai -400101, Maharashtra, India

Abstract

This paper describes the comparative study of Truncated DCT-SVD and DWT-SVD. In this paper we propose two different approaches to compute the feature vector for content based image retrieval (CBIR) system. SVD feature of successively truncated DCT image and DWT decomposed image computed for grayscale image, RGB and YCbCrcolor image. Truncated DCT and DWT decomposition SVD features of the image computed up to fifth level to compare the performance. Proposed methods incorporate with the multidimensional features vector computed by using SVD of low frequency coefficients of DCT and DWT of image. Similarity between the query image and database image measured here by using simple Euclidean distance and Bray Curtis Distance. The overall average precision and average recall crossover point of each image category. Proposed methods are tested on the augmented image database.

Peer-review under responsibility of the Organizing Committee of ICCCV 2016

***Keywords*:CBIR; DCT; DWT; SVD; DCT-SVD;DWT-SVD, Overall Average Precision and Recall crossover point.**

1. Introduction

CBIR or Content Based Image Retrieval is the retrieval of images based on visual features such as color, texture and shape. Reasons for its development are that in many large image databases, traditional methods of image indexing have proven to be insufficient, laborious, and extremely time consuming. These old methods of image indexing, ranging from storing an image in the database and associating it with a keyword or number, to associating it with a categorized description, have become obsolete. This is not CBIR.

[1] Email id :hsainst@gmail.com ,[2] Email id : ravi.patil212@gmail.com,[3] Email id :sss.sangle@gmail.com,[3] Email id : bnemade@gmail.com

Peer-review under responsibility of the Organizing Committee of ICCCV 2016
doi:10.1016/j.procs.2016.03.073

In CBIR, each image that is stored in the database has its features extracted and compared to the features of the query image. It involves two steps:

- Feature Extraction: The first step in the process is extracting image features to a distinguishable extent.
- Matching: The second step involves matching these features to yield a result that is visually similar.

Content-based image retrieval relies on the characterization of primitive features such as color, shape, and texture that can be automatically extracted from the images themselves. Commercial CBIR systems in use include IBM's Query by Image Content (QBIC) described first by Flickneret al, (1995), Virages' VIR Image Engine (Gupta et al, 1996), and Excalibur's Image Retrieval Ware. On the Web, CBIR image retrieval systems include Web SEEK (Smith & Chang, 1997), Informedia, and Photo book among others. Comparison of two image descriptors for image retrieval is presented in this paper. Content based image retrieval is well known research area in the information management in which large number of feature vectors are generated but still no satisfying general solutions exist. This area has been active and popular research area in information management because over the past few decades due to its wide and potential applications, such as journalism, medicine, and private life, requiring new ways of accessing images. For example, medical doctors have to access large amounts of MIR images daily, home-users often have image databases of thousands of family photos,and journalists also need to search for images by various criteria (Markkula and Sormunen 1998; Armitage and Enser 1997)[1]. In the past, several CBIR systems have been proposed and all these systems have one thing in common: images are represented as descriptors that describe the properties of images such as color, texture and shape. Color image descriptor is very easy and efficient in image retrieval system, because it is less complex and compact. Texture and shape image retrieval complexity dependent on type of algorithms are used to compute the feature vector.

2. History of Existing Techniques

Color image retrieval isretrieval of images based on visual descriptor such as color, texture and shape. Sometimes color descriptor fail to retrieve the relevant image from the database. Texture refers to the visual patterns that have properties of homogeneity that do not result from the presence of only a single color or intensity. It is an innate property of virtually all surfaces, including clouds, trees, bricks, hair, and fabric. It contains important information about the structural arrangement of surfaces and their relationship to the surrounding environment. The Discrete Cosine Transform is a member of a family of real valued discrete sinusoidal orthogonal transforms. It consists of a set of basis vectors that are sampled cosine functions[2]. The DCT has very strong energy compaction capability so it is widely used in image compression. Hee-Jung Bae et al. propose a new content-based image retrieval method using texture information is proposed[3]. For efficient image retrieval, it extracts low-level image texture features as content retrieval index. In Kekre et al. presents 32 novel image retrieval techniques using the feature vectors obtained by applying Discrete Cosine Transform on row mean and column mean of full image, four fragments, sixteen fragments and 64 fragments of image using gray scale and RGB image[4,5].

The Discrete Wavelet Transform (DWT) [6, 7] has a large number of applications in science engineering mathematics and computer science due to its link to digital filtering. It is used for signal coding to represent a discrete signal in a more redundant form often as a preconditioning for data compression. Wavelet decomposition separates the high frequency and low frequency components from given image[7, 8]. In the early 1990s, after the wavelet transform was introduced and its theoretical framework was established, many researchers began to study the use of the wavelet transform in texture representation..

As explained in D. Kahaner et al. SVD[11] is widely used in various fields like image compression, image water marking, image coding and introduced in our thesis specially for feature extraction, functioning similarly as in [85, 86] using singular values as the robust feature for its geometrical invariance. In [12] G. He et al. discussed the reasonability that singular values be used for similarity evaluation. In [12] and [13] relied on SVD for the face identification applications which seldom involve complex signal processing modifications except geometrical transformations. Thomas J. Peters et al. [14,15] proposed microarray image compression using singular value decomposition (SVD), a well-known information compaction method. Although they prove the SVD algorithm produces significant compression results, modifications may lead to further improvements.

This paper presents a new feature vector based on truncated DCT-SVD and DWT-SVD, which is compare size of the feature vector and retrieval performance. The paper is organized as follows: problem statement and proposed idea and feature vector extraction in Sections 3 and 4, respectively. Section 5 explains the performance parameter Sections6 explains the experimental resultdiscussion and Section 7conclusions are presented.

3. Problem Statement

Descriptor size of an image is variable it should be compact. It requires more memory to store these descriptors.Above method fractional DCT coefficients as a feature vector require more space to store even as the feature vector size goes on increasing the retrieval result is better in terms of overall precision and recall crossover point. Wavelet based features were first introduced in[16] Jacobs et al. selects sixty four largest Haar wavelet coefficients in each of the 3 color band and stores them in feature vector as +1 or -1 along with their position in the transformation matrix. Low frequency coefficients tend to be more dominant than those of the high frequency coefficients and this makes this algorithm ineffective for images with sharp color changes. Still they cannot explicitly address the feature vector size problem. More recently,proposed a novel fractional order singular value decomposition representation method for face recognition. Their motivation is each grayscale face image can be decomposed into a composition of a set of bases by the well-known singularvalue decomposition (SVD) technique.

4. Proposed Idea

To mitigate above problems in we propose two methods which are listed below

4.1. Singular Values of Successive Truncated DCT Image

4.2. Feature Vector Extraction using DWT-SVD Domain

- *Singular Values of Successive Truncated DCT Image*

When 2D-DCT applied on the NXN image then NXN DCT transformed coefficients are obtained. In these DCT coefficients upper left first coefficient is DC coefficient which represents average luminance.

Fig 1.Given image and its first level, second level, third level DCT truncation.

Remaining coefficients are the AC coefficients which represent the intensity change in the pixel value. The low frequency coefficients concentrate at the upper left corner of the transformed plane and high frequency coefficient are the remaining part. Take DCT of given image having size NXN and then truncate low frequency coefficients. Apply the inverse DCT on the low frequency coefficients. Image having size N/2XN/2 will obtain. Fig.1 show the original gray scale rose image and its DCT truncated images up to third level.Take SVD of truncated DCT image. Singular value having size 1 x N/2 obtained. To find out impact of different singular value, truncate these singular values in N/2,N/4,N/8,N/16……so on as shown in Fig.2 These truncated singular values used for feature vector formation. At first level singular values are truncated in five groups that mean five feature vectors are formed at first level, four at second level, three at third level, two at fourth level and one at fifth level. Image used for this process having size 256x256. So for the given image as listed in table 1 each level singular values are divided. For example at 1st level take first 8 singular values as a feature, first 16 singular values, first 32 singular values, 64 singular values, 128 singular values for feature vector. So comparison can be take place for performance at each level of successive truncated DCT.

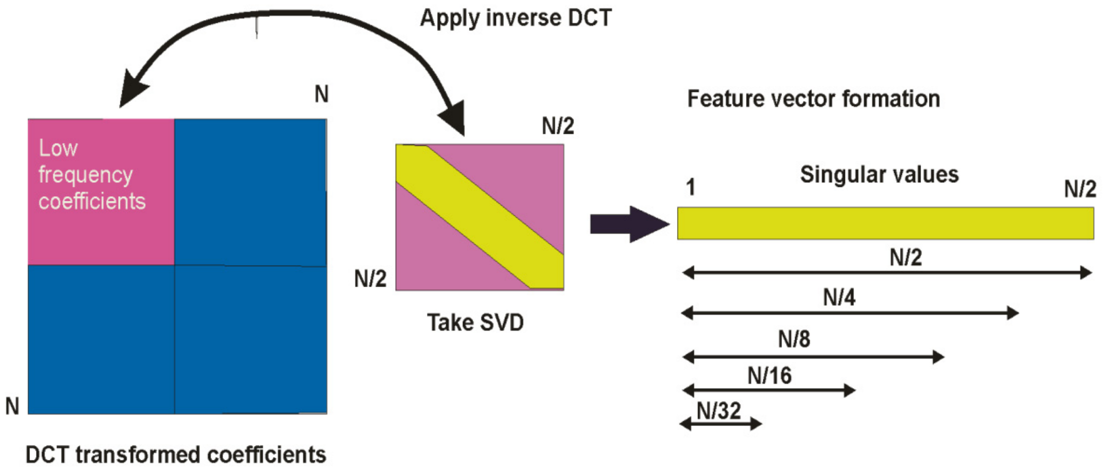

Fig 2.Feature vector formation for the truncated DCT image SVD.

Table 1.Division of singular values at each level of DCT truncation.

DCT Truncation	Ranges of Singular Value coefficients as a feature vector
1st level	8,16,32,64,128
2nd level	8,16,32,64
3rd level	8,16,32
4th level	8,16
5th level	8

- *Feature Vector Extraction using DWT-SVD Domain*

In this proposed approach singular values of Discrete Wavelet Transform (haar wavelet) sub bands coefficients are used as a feature vector. DWT decompose an image into pyramid structure of sub image with various resolutions. Feature vector is computed using singular values of decomposed image [17,18]. 1st level decomposition of given image gives LL, LH, HL and HH sub images. Take SVD of each sub band image. For feature vector concatenate each sub band singular values together. So at first level 4x N/2 size feature vector is obtained. This is the 1st level feature vector computation process using DWT as shown in Fig 3. Again go for the second level decomposition compute the singular values of 2nd level sub band images and concatenate them to make a feature vector. Similarly up to 5th level of decomposition find out the feature vector using SVD. Thus the feature vector database for DWT is generated as DWT-SVD. Feature vector size increases from 1st level of decomposition to the 5th level.

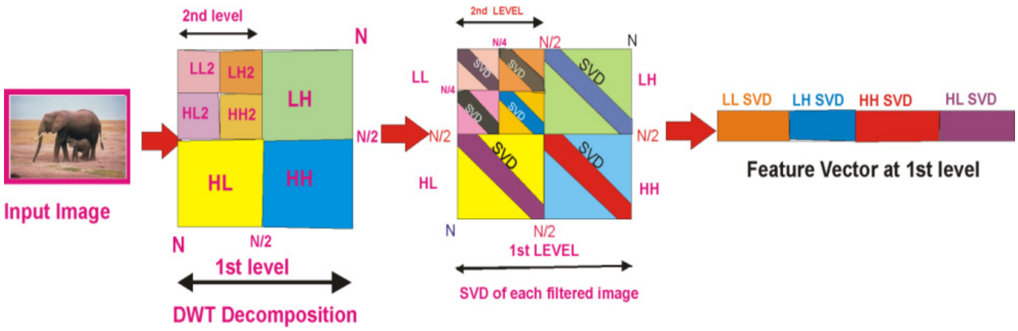

Fig. 3. Feature vector calculation for DWT-SVD.

5. Performance measurement

Testing the effectiveness of the CBIR is about testing how well the CBIR can retrieve similar images to the query image and how well the CBIR prevents the return results that are not relevant to the source at all in the user point of view. A sample query image must be selected from one of the image category in the database. When the search engine is run and the result images are returned, the user needs to count how many images are returned and how many of the returned images are similar to the query image. Determining whether or not two images are similar is purely up to the user's perception. After images are retrieved, the CBIR approach effectiveness needs to be determined. To achieve this, two evaluation measures are used. This evaluation diagrammatically represented in the fig.4.The first measure is called Recall. It is a measure of the ability of a system to present all relevant items. The second measure is called Precision. It is a measure of the ability of a system to present only relevant items.

The number of relevant items retrieved is the number of the returned images that are similar to the query image in this case. The number of relevant items in collection is the number of images that are in the same particular category with the query image. The total number of items retrieved is the number of images that are returned by the search engine as shown in Fig.4. Precision stipulate the accuracy of the result of CBIR technique and recall specifies the completeness. Higher value of precision and recall indicated better image retrieval.

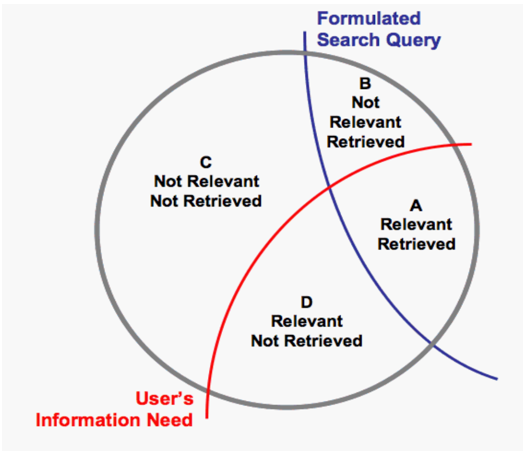

Fig 4. Performance of CBIR

Usually a threshold will be determined on trial and error basis which is used to retrieve the images similar to query having distance less than pre-determined threshold. As this trial and error method is time consuming process and never comes up with fix or ideal threshold value, what we do here for retrieval is, we sort the distances in ascending order and selecting first N number images to retrieve the images relevant to the query image out of it. N is the total number of images of that particular category in the database images. This generates the cross over points [5,6,19]of precision and recall values for that particular category our results. This means that as average precision goes on decreasing recall goes on increasing with respect to number of images retrieved. A point at which average precision and recall values are same as shown in Fig. 5.The crossover point of the precision and recall play very important role in the performance of analysis of the image retrieval method which act as a performance measure of CBIR. This crossover point value is the value where precision equal to the recall. In ideal situation the height of precision and recall cross over point should be at the value one.

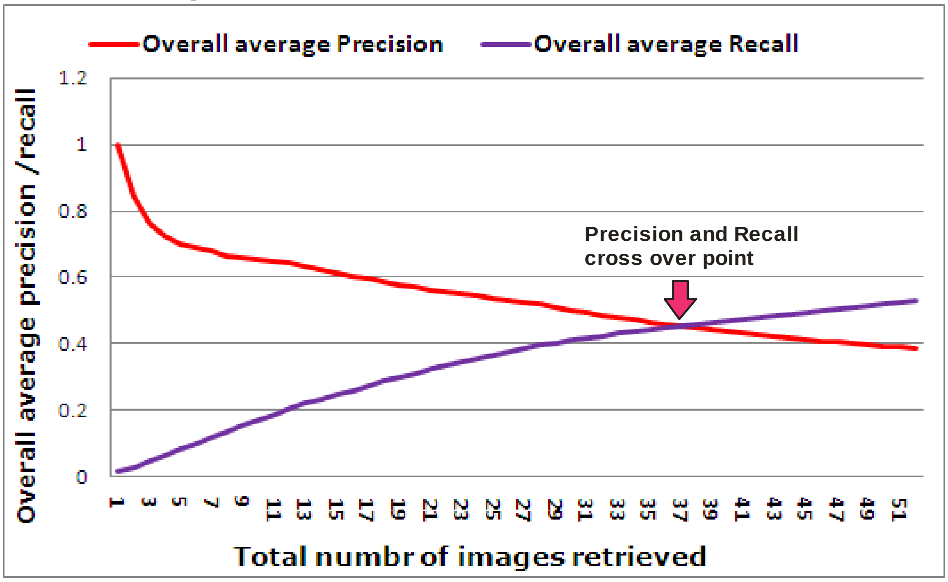

Fig 5.Average precision and recall crossover point

This means all the retrieved images are relevant and all relevant images from the database are retrieved. Always, performance of the retrieval technique compared with ideal situation. The height of precision and recall crossover point gives idea about how much the proposed technique is deviating from the ideal one. More the height better is the technique.

5.1. Minkowski-form Distance(ED)

If each dimension of image feature vector is independent of each other and is of equal importance, the Minkowski-form distance[1] is appropriate for calculating the distance between two images. This distance is defined as,

$$D(I,J) = (\sum_i |f_i(I) - f_i(J)|^P)^{1/P} \tag{1}$$

When p=1, 2, and ∞, D (I, J) is the L1, L2 (also called Euclidean distance), and L∞ distance respectively. Minkowski-form distance is the most widely used metric for image retrieval. For instance, MARS system used Euclidean distance to compute the similarity between texture features; Netra used Euclidean distance for color and shape feature, and L1 distance for texture feature; Blobworld used Euclidean distance for texture and shape feature. In addition, Voorhees and Poggio used L∞ distance to compute the similarity between texture images.

5.2. Bray Curtis Distance(BCD)

$$Bd(Q,I) = \frac{\sum_{k=1}^{n} \left| H_{Qk} - H_{Ik} \right|}{\sum_{k=1}^{n} (H_{Qk} - H_{Ik})} \tag{2}$$

Where
n-Total no of component in feature vector.
Q-Query image
I- Database image.
HQk-Feature vector query image.
HIk-Feature vector for database image.

Bray Curtis distance sometimes is also called Sorensen distance is a normalization method that common used in botany, ecology and environmental field. It views the space as grid similar to the city block distance. The Bray Curtis distance has a nice property that if all coordinates is positive, its value is between zero and one. Zero bray Curtis represent exact similar coordinate. If both objects are in the zero coordinates, the Bray Curtis distance is undefined. The normalization is done using absolute difference divided by the summation.

6. Experimental result discussion

Proposed approach tested on the augmented Wang database which includes 1200 variable size images spread over 15 different category. The implementation of the discussed CBIR techniques is done in MATLAB 7.0 using a computer with Intel Core 2 Duo Processor T8100 (2.1GHz) and 2 GB RAM. The categories and distribution of the images is shown in table 2.

Table 2. Category-Wise Distribution of Image Database

Name of Category	No. of Images	Name of Category	No. of Images	Name of Category	No. of Images
Motorbikes	100	Beaches	100	Buses	100
Elephants	100	Flowers	100	Tribal Peoples	68
Flying Birds	63	Flower lawn	48	Butterfly Scenery	52
Dinosaurs	100	Mountains	62	Guitar	59

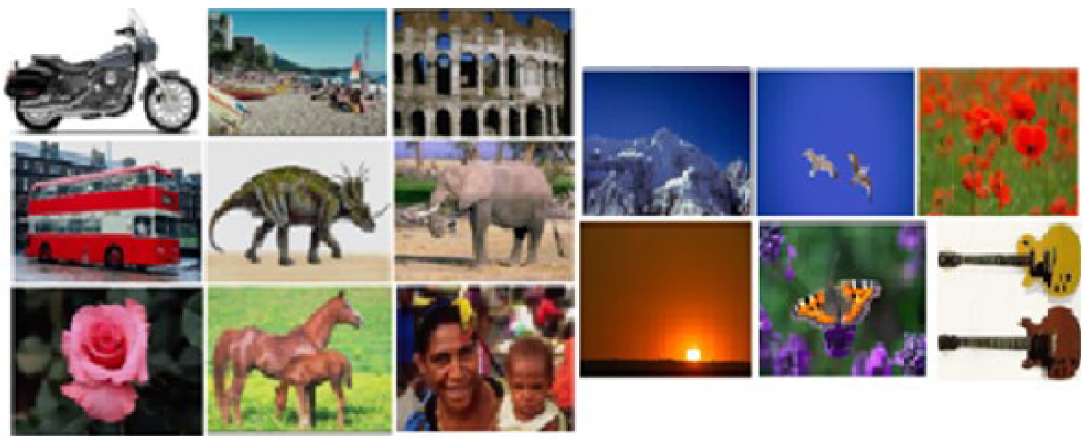

Fig. 6 Sample images from database.

Sample images are shown in Fig.6.Query images (5 from each class) are fired on the database and similarity matching is done by using Euclidean distance and Bray Curtis distance measure. The average precision and average recall values are computed by grouping the number of retrieved images sorted according to ascending distances with the query image. Average precision and recall crossover points performance of proposed methods plotted against the each class of image with the two similarity measures as shown in Fig.7.shows performance plots only for the level which have highest cross over point compared to the above approaches.

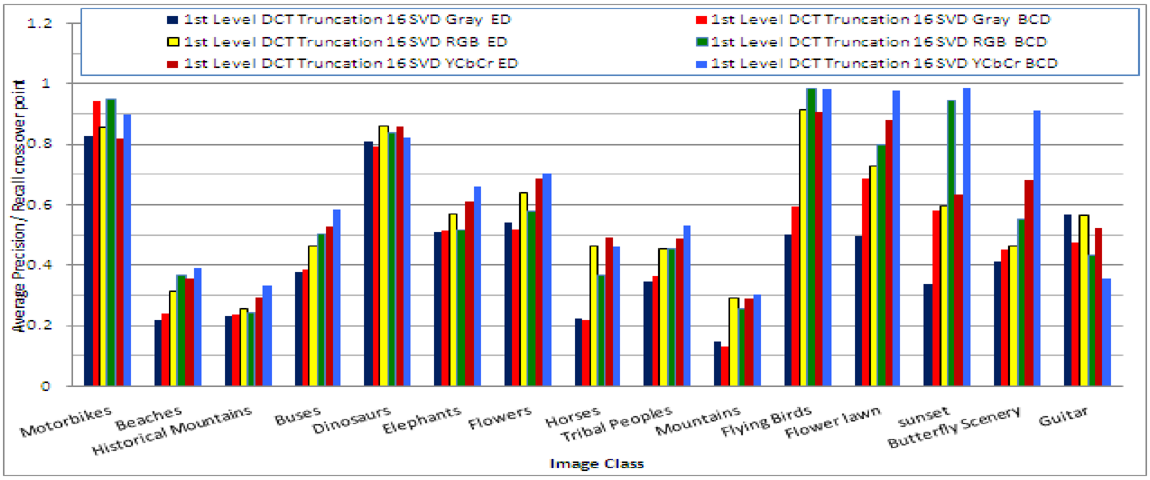

Fig.7. Comparisons of average precision and recall crossover point for 1st level 16 singular values of truncated DCT image as a feature vector of all class of images using Euclidean distance (ED) and Bray Curtis distance similarity.

That means at 1st level 16singular values performance. It is seen on the plots that the motorbike, dinosaurs, flying birds, flower, flower lawn, sunset, butterfly sensory image classes performing better than the other classes in terms of individual average precision and recall cross over points and it is above approximately 45%. Flying birds and flower lawn cross over points performances approaches towards 100% using YCbCrcolor image and RGB color image with Bray Curtis distance similarity.

Table 3. Overall average precision and recall cross over points for SVD of truncated DCT.

DCT Truncation level	Color Planes	Overall APRCP for 8 SV		Overall APRCP for 16 SV		Overall APRCP for 32 SV		Overall APRCP for 64 SV		Overall APRCP for 128 SV	
		ED	BCD	ED	BCD	ED	BCD	ED	BCD	ED	BCD
1st Level	Gray	0.425	0.454	0.438	0.477	0.445	0.490	0.449	0.460	0.452	0.505
	RGB	0.552	0.580	0.562	0.585	0.564	0.576	0.565	0.554	0.569	0.582
	YCbCr	0.591	0.652	0.604	0.662	0.609	0.658	0.609	0.646	0.614	0.658
2nd Level	Gray	0.422	0.453	0.433	0.472	0.439	0.489	0.433	0.443	--	--
	RGB	0.549	0.567	0.558	0.561	0.559	0.544	0.559	0.589	--	--
	YCbCr	0.591	0.644	0.599	0.642	0.603	0.632	0.606	0.665	--	--

Table 3 shows the overall average precision and recall cross over point for the SVD of truncated DCT image at each level using YCbCrcolor image, gray scale image and RGB color plane. At truncated DCT 1st level 8,16,32 and 128 singular value as a feature vector performances and at the second level truncated DCT 64 singular values as a feature vector performances are better than the other levels. It is noted that the overall average precision and recall cross over point(APRCP) for 64 singular values(SV) of 2nd level truncated DCT image as feature vector using YCbCrcolor image with Bray Curtis distance similarity is higher and it is above 66%.Other feature vectors performances are above 65% for 8 singular values ,above 66% for 16 singular values above 65% for 32 singular value and above 65% for 128 singular values as a feature vector of truncated DCT first level using YCbCrcolorimage with Bray Curtis distance similarity.

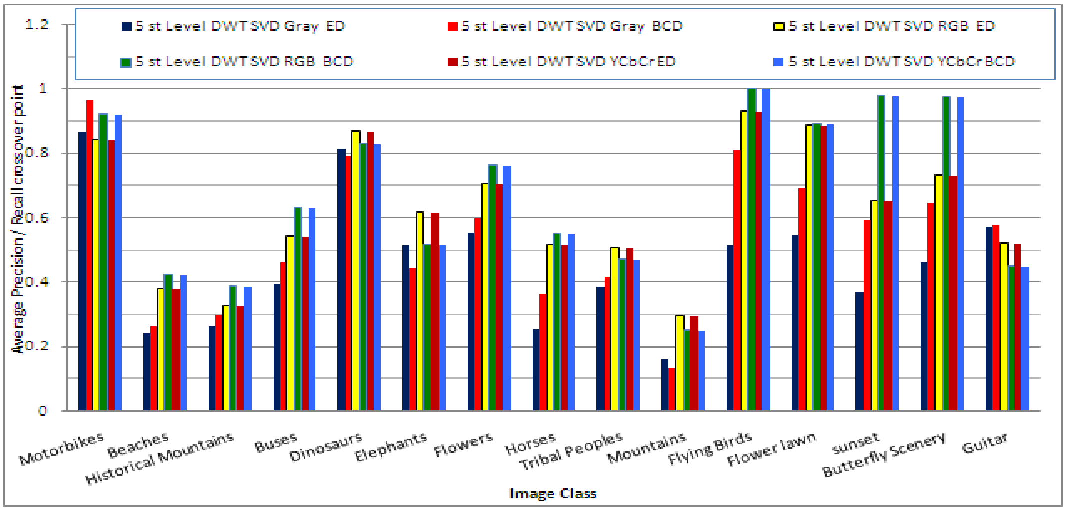

Fig.8 Comparisons of average precision and recall crossover points for singular values of sub bands of 5^{th} level DWT decomposition as a feature vector of all classes of images using Euclidean distance (ED) and Bray Curtis distance (BCD) similarity.

Fig.8 shows the average precision and recall cross over points plotted against the image classes for proposed image retrieval technique. In Fig.6 show the plots for 5th level DWT decomposition–SVD cross over points for all image classes. Here flying birds, sunset and butterfly sensory image classes performances using RGB, YCbCrcolor image with Bray Curtis distance similarity are higher and it reaches towards 100%. Motorbike, Dinosaur and flower image classes performances are above 60% for both the similarities.

Table 4. Overall average precision and recall cross over points for DWT-SVD domain feature.

Feature vector computation Method	Color Planes	Overall average precision and recall cross over point	
		ED	BCD
1st level DWT Decomposition SVD	Gray	0.425	0.454
	RGB	0.552	0.580
	YCbCr	0.591	0.652
2nd level DWT Decomposition SVD	Gray	0.422	0.453
	RGB	0.549	0.567
	YCbCr	0.591	0.644
3^{rd} level DWT Decomposition SVD	Gray	0.4562	0.5469
	RGB	0.5756	0.6171
	YCbCr	0.6169	0.6794
4^{th} level DWT Decomposition SVD	Gray	0.4518	0.5482
	RGB	0.5726	0.6209
	YCbCr	0.6158	0.6817
5^{th} level DWT Decomposition SVD	Gray	0.4481	0.5484
	RGB	0.5693	0.6235
	YCbCr	0.6130	0.6843

Table 4 shows the overall average precision and recall cross over points for DWT-SVD domain approach. It is seen that highest precision and recall cross over point is above 68% for the 5th level decomposition using YCbCrcolor image. YCbCrcolor image performance for both the similarities are above 61% at each decomposition level. For RGB color image it is above 64% and for gray scale images it is above 54% with Bray Curtis distance similarity. By using Bray Curtis distance similarity precision and recall cross over point's increases from 1st level to 5th level DWT decomposition. But Euclidean distance similarity crossover point not consistently increases but it is fluctuating.

7. Conclusion

The new ideas of computing feature vector using DWT decomposition image SVD and DCT truncated image SVD up to 5th level for content based image retrieval is proposed. The proposed methods are tested on gray scale image, RGB color image and YCbCr color image. Performance in terms of overall average precision and overall average recall cross over points of proposed methods with respect to each level and for each color plane is compared. We found that the proposed methods performance is acceptable .It is on an average 50% DCT-SVD and 54 % DWT-SVD for gray scale image using ED similarity measure, 62% DWT-SVD method and 59% for DCT-SVD method for RGB color image and 68% DWT-SVD method and 66% for DCT-SVD method for YCbCr color image for ED similarity measure. In all the proposed 8,16 SV DCT- SVD algorithms perform better for YCbCr color image using BCD similarity.

References

1. T.Deselaers,D.Keysers,H.Ney,:Features for image retrieval: an experimental comparison, springer,Information Retrieval (2008) 11:77–107 DOI 10.1007/s10791-007-9039-3.

2.N.Ahmed, T. Natarajan, and K.R.Rao:Discrete cosine transform ,IEEE Trans. on Comput.,vol.23,pp.90-93,1974.

3.Hee-Jung Bae, Sung-Hwan Jung :Image Retrieval Using Texture Based on DCT,International Conference on Information, Communications and Signal Processing ICICS '97 Singapore,7803-3676-3/1997 IEEE ,9-12 September 1997.

4.H.B. Kekre, Sudeep D. Thepade, ArchanaAthawale, AnantShaha,PrathmeshVerlekar, SurajShirke:Grayscale Image Retrieval using DCT on Row mean, Column mean and Combination, TechnoPath :Journal of Science Technology and Management, Volume 2, Issue 1, January 2010. 5. H.B. Kekre, Sudeep D. Thepade, ArchanaAthawale, AnantShaha,PrathmeshVerlekar, SurajShirke:Grayscale Image Retrieval using DCT on Row mean, Column mean and Combination, TechnoPath :Journal of Science Technology and Management, Volume 2, Issue 1, January 2010.

6.T. Changand C.J. Kuo:Texture analysis and classification with tree-structured wavelet transform, IEEE Transaction Image Proc. 2(4), 429–441, 1993.

7. M. H. Gross, R. Koch, L. Lippert, and A. Dreger:Multiscale image texture analysis in wavelet spaces, in Proc. IEEE Int. Conf. on Image Proc., 1994.

8. A. Laine and J. Fan:Texture classification by wavelet packet signatures, IEEE Trans. Patt. Recog. andMach.Intell. 15(11), 1993, 1186–1191.

9. J. R. Smith and S. F. Chang:Transform features for texture classification and discrimination in large image databases", in Proc. IEEE Int. Conf. on Image Proc., 1994.

10. J. R. Smith and S.F. Chang:Automated binary texture feature sets for image retrieval, in Proc. ICASSP-96, Atlanta, GA, 1996.6.wikipedia.org,"AdvancedEncryption Standard", https://en.wikipedia.org/wiki/Advanced_Encryption_Standard.

11. D. Zhang, A. Wong, M. Indrawan, G. Lu:Content based image retrieval using Gabor texture features, in: Proceedings of 1st IEEE Pacific Rim Conference on Multimedia (PCM'00), 2000, pp. 1139–1142.

12. D.Kahaner, C.Moler and S. Nash:Numerical methods and software" ,(New Jersey: Prentice-Hall, Inc,(1989).

13. G. He and J. Gan:A method for singular value feature extraction of face image, Proc. of 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing, Hong Kong, October 2004.

14. Y. Wang, T. Tan, Y. Zhu:Face verification based on singular value decomposition and radial basis function neural network, Proc. 4th. Asian Conference on Computer Vision, pp.432-436, Taipei, Taiwan, 2000.

15. Thomas J. Peters, RenataSmolíková-Wachowiak:Microarray Image Compression Using a Variation of Singulqr Value Decomposition, Proceedings of the 29th Annual International Conference of the IEEE EMBS CitéInternationale, Lyon, France August 23-26, 2007.

16.C.E. Jacobs, A. Finkelstein, and D.H. Salesin:FastMultiresolution Image Querying" ,Proc. SIGGRAPH 95, 1995.

17. H. Ochoa, K.R. Rao:A Hybrid DWT-SVD Image-Coding System (HDWTSVD) for Monochromatic Images,WSEAS Transaction n Circuit and Systems, vol. 4, pp. 419 – 424, Dec. 2005.

18. Ochoa, H.Rao, K.R.A :modified hybrid DWT-SVD image coding system (HDWTSVD) for monochromatic images,4th EURASIP Conference on Video/Image Processing and Multimedia Communications,2003.July 2003 Volume:1,pp 323-328.

19. H.B.Kekre, Sudeep D. Thepade,Image Retrieval using Color-Texture Features Extracted from WalshletPyramid,ICGST - GVIP Journal, ISSN: 1687-398X, Volume 10, Issue 1, February 2010.